ISSN 2229-5518

Harmonic Analysis of Sunspot Numbers

Iwok, Iberedem Aniefiok

Department of Mathematics/Statistics,

University of Port Harcourt, P.M.B. 5323 Port Harcourt,

Rivers State, Nigeria.

e-mail: ibywok@yahoo.com

Abstract

This work modeled the sunspot numbers using harmonic analysis. A period of 10.8 years identified by Iwok (2011) was used in building the model. It was observed that the fitted sinusoid accounted for 71.7% of the variance of the observed time series X_t. The fitted model was also found to be adequate and can be used for predicting the future of sunspot numbers.

Keywords: Amplitude, Period, White noise, Sinusoid, Phase angle.

Introduction/Review

Sunspots are dark spots within the outermost layer of the sun called the photosphere. The photosphere is about 400km deep and provides most of our solar radiation. Sunspots were observed in the far East over 2000 years ago but was examined more intensively in Europe after the invention of telescopes in the 17th century. Since the mid eighteen century, an observatory in Switzerland has been recording monthly, a measure of the number of sunspots, for reasons of their own and presumably at considerable risk to the astronomer's eyesight, thus providing one of the longest time series available for analysis (Granger, 1989).

Sharaf and Amr (2009) noted that sunspots are a key factor and a good usable indicator of solar activity for sun-weather relationship. In their analysis, they pointed out that an obvious condition must be met before sunspot numbers can be used to predict changes in weather and climate. Sharaf and Amr (2009) carried out a data analysis of the sunspot numbers using annual mean and autocorrelations. Confidence intervals were constructed and the results were supported graphically and computationally by many tests.

The well known 11 year cycle of the number of sunspots was first demonstrated by Henrich Schwab in 1843.

Richard (2005) confirmed the existence of cycles in sunspot numbers using early historical methods, and this was calculated to be 11.1 years.

Iwok (2011) identified the cycles in sunspot numbers using percentage variance technique and was found to be 10.8 years. By this calculation, it means that sunspot numbers rise and fall approximately every 10.8 years.

Due to the enormous importance of sunspot study to man and his environment, researchers have fitted several models to sunspot numbers over the years, ranging from the simple linear approach to non linear methods.

Tsirulnik et al (1997) used a method of non linear spectral analysis named by the method of the global minimum to find periodicities in the annual wolf sunspot numbers for the 295 year time period (1700-1995). It was discovered that the normalized mean square error of the model fitted well to the data and the magnetic cycle were found to be non stationary with a mean period of 22.3 years.

Granger (1989) fitted a model of the form:

$$X_t = ACos(\omega t + \theta) + \varepsilon_t \qquad (1)$$

to the sunspot series and a worthwhile fit was obtained but the estimated residuals were observed to be too smooth to be white noise.

Similarly Kenedy et al (2010) modeled the sunspot numbers using Fourier analysis. It was discovered that despite the complexity of the method, the white noise ε_t failed the assumption of being independently normal and identically distributed random variables.

This research, however, attributes the failures of the afformentioned works to the use of inappropriate models to cater for the cycles and hence the need for modifications and extensions. It is therefore the intent of this work to propose and fit a model that gives ε_t as a purely random process.

Method of Analysis

This research uses harmonic form of time series as a type of regression analysis. By definition, a time series is a series of observation made sequentially in time. The yearly recordings of the sunspot numbers (*see table 2*) are time series and since the raw plot (*see figure 1*) reveals a cyclic nature, this work shall employ the harmonic technique in analyzing the sunspot series. The harmonic analysis involves the creation of two predictor variables: one time series that represents a sine of period τ and another that represents a cosine of period τ.

We recall that a simple bivariate linear regression equation that describes any linear trend of a time series is:

$$X_t = a + bt \qquad (2)$$

where t is the time index and a and b are parameters to be estimated by ordinary least squares (OLS) method. It should be noted that the above linear equation applies when there is a clear linear trend in the time series X_t. In most cases, however, the time series tends to show regular, sinusoidal cycles (*see figure 1*) thereby calling for a more complex model with trigonometric functions.

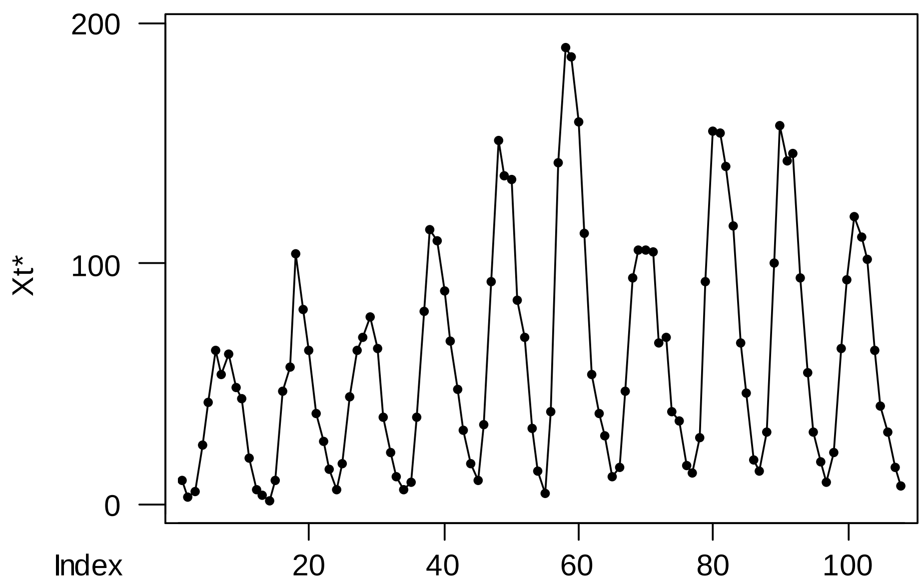


Figure 1: Raw data plot of the Sunspot Numbers

Let us replace t in equation (2) with a trigonometric function of t that represents a sinusoid and add the parameters that allow us to estimate the phase and amplitude that best fit the observed time series X_t. Such representation will form the basis for harmonic analysis since it involves fitting a sinusoid to a time series. Similar to Granger (1989), we express the form:

$$X_t = \mu + ACos(\omega t + \emptyset) + \varepsilon_t \quad (3)$$

where μ is the mean added to give the level of the time series X_t.

A is the amplitude of the wave form, $\emptyset$ is the phase or location of peaks relative to time zero,

ε_t represent residuals expected to give a white noise process if the fitted model is appropriate.

A white noise is a sequence of random variables from a fixed distribution usually assumed normal and having mean zero and constant variance (Box and Jenkins, 1970).

$\omega = \frac{2\pi}{\tau}$ is the corresponding frequency in radians.

Alternatively, we can represent a more general case of the sinusoid by adding the sine component as follows:

$$X_t = \mu + ACos\omega t + BSin\omega t + \varepsilon_t \qquad (4)$$

Here, $\emptyset$ is relaxed to ease computation of parameter estimates. To obtain $\emptyset$, we can set $tan\emptyset = \frac{-B}{A}$ (Bloomfield, 1976). It is clearly seen from equation (4) that by varying A and B, it is possible to generate all possible sinusoids of period τ.

Provided, the period τ is known, all the parameters $(\mu, A \; and \; B)$ of equation (4) can be estimated using the OLS regression method. Hence predictions of the future can be made at any time t.

Data Analysis and Interpretation

The data used for the study was the annual averaged sunspot numbers (1900-2007). The analysis was performed with the aid of software called Minitab. A period of 10.8 years identified by Iwok (2011) is used in the OLS harmonic modeling of the sunspot numbers. The minitab output of the analysis is displayed in table 1 below:

Table 1: Minitab output

Regression Analysis

The regression equation is
Xt* = 60.6 - 44.4 CosWt* + 18.7 SinWt*

Predictor	Coef	StDev	T	P
Constant	60.619	3.213	18.87	0.000
CosWt*	-44.406	4.543	-9.77	0.000
SinWt*	18.661	4.543	4.11	0.000

S = 33.39   R-Sq = 71.7%   R-Sq(adj) = 70.8%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	2	125284	62642	56.20	0.000
Error	105	117035	1115		
Total	107	242319			

Using equation (4), the estimated harmonic equation that describes the cyclic pattern is:

$$X_t = 60.6 - 44.4Cos\omega t + 18.7Cos\omega t \qquad (5)$$

with $\omega = \frac{2\pi}{10.8}$.

The estimate of the overall mean of X_t is the intercept of equation (5) which is 60.6. This indicates the level about which X_t fluctuates. R^2 = 71.7% shows that a 10.8 years cycle account for 71.7% of the variance in X_t. Such high percentage is an indication of good fit of the model (5).

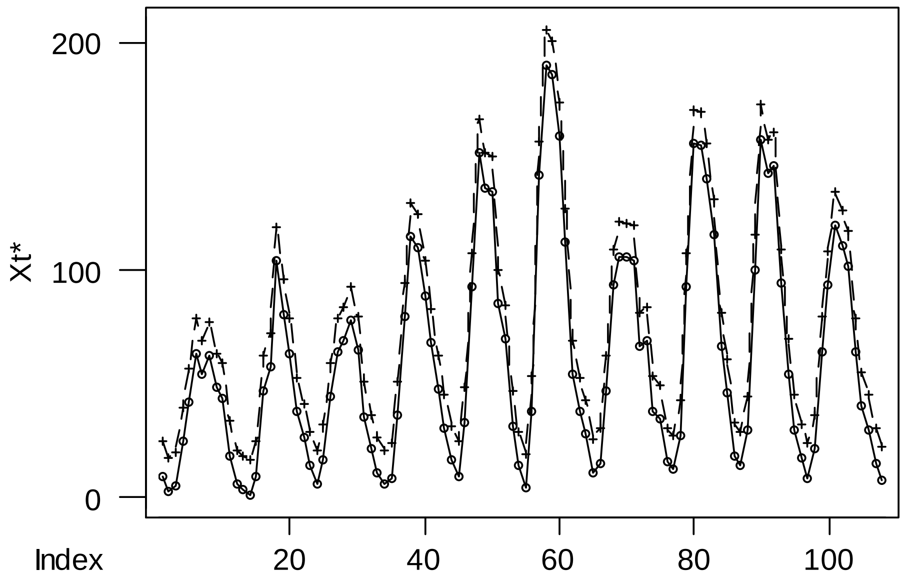
The amplitude, $A^* = [A^2 + B^2]^{\frac{1}{2}} = [(-44.4)^2 + (18.7)^2]^{\frac{1}{2}} = 48.8$. That is, the peaks and troughs of the sinusoids are about 48.8 points above and below 60.6.

Actual and Estimate Plot

The estimates of the sunspot numbers were computed from the model:

$$X_t = 60.6 - 44.4Cos\left(\frac{2\pi}{10.8}t\right) + 18.7Sin\left(\frac{2\pi}{10.8}t\right) \quad (6)$$

The plot of the estimated values was superimposed on the plot of the raw data to give figure 2 below:



Key: o(actual plot) and +(estimate plot)

Figure 2: Actual and Estimate Plot of the Sunspot numbers X_t

The plot in figure 2 reveals a strong agreement between the real and estimated values of X_t, thus indicating a good fit of the model.

Residual Plot

Residual ε_t is obtained from equation (5) as:

$$\varepsilon_t = X_t - [60.6 - 44.4Cos\omega t + 18.7Sin\omega t] \quad (7)$$

After fitting the model, the residual obtained was plotted as shown in figure 3 and its variance was 21.3.

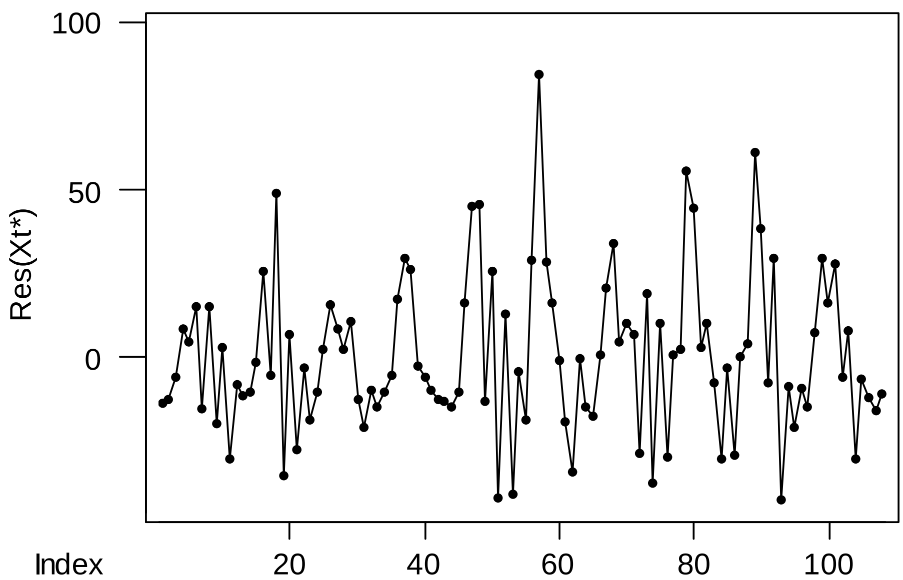


Figure 3: Residual Plot of X_t

It is clearly seen from figure 3 that the residual plot gives a white noise process. That is,

$\varepsilon_t \sim NIID(0, \sigma^2)$.

Summary/Conclusion

The study of sunspot numbers has engaged the attention of many researchers due to its applicability in climatology especially at this period of combat with global warming. This work is another contribution made to generate a model for forecasting sunspot numbers.

The sunspot cycle was identified to be 10.8 years (Iwok, 2011) using the percentage variance technique. This work used the identified cycle to establish a forecasting model based on harmonic analysis as a form of regression. Using the sunspot data in table 2 and the residual checks, the model is found to be adequate and can be used for predicting the future of sunspot numbers.

References

Blooomfield, P. (1976). Fourier analysis of time series: An introduction. New York, Wiley

Box, G.E.P. and Jenkins, G.M.(1970). Time series analysis, forecasting and control. San Francisco Holden Day. New York.

Granger, C.W.J. (1989). Forcasting in business and economics(second edition). Academic press inc; Harcourt brace. Jovanovich publishers. Boston, New York.

Iwok, I. A. (2011). Identification of Cycles in sunspot numbers using percentage variance technique. American Journal of Scientific and Industrial research. Vol.2, No. 4, pp. 488-490. © 2011, science Hub, http://www.scihub.org/AJSIR. ISSN: 2153-649. DOI:10.5251/ajsir.2011.2.4.488.490.

Kenedy, M. et al (2010). Fourier analysis of sunspot numbers. International Journal of Science and Technology. Vol. 2, No. 6, pp. 51-62.

Richard, H.T.(2005). Time series analysis of the sunspot cycle. Indian Journal of applied statistics. Vol.3, No.15, pp 46-55.

Sharaf, M.A. and Amr, A.S.(2009). Data analysis of the annual mean of the sunspot numbers. JKAU Science. Vol. 21, No. 2, pp. 371-379. DOI: 104197/Sci.21-2,14.

Tsirulnik, L. B. et al (1997). Advances in space research. Vol.20, Issue 12, pp. 2365-2368.

ISSN 2229-5518

Table 2 : Annual Averaged Wolfer's Sunspot Numbers X_t (1900-2007)

Courtesy: *Solar Index Data Centre*

YEAR	VALUE	YEAR	VALUE	YEAR	VALUE	YEAR	VALUE	YEAR	VALUE	YEAR	VALUE
1900	9.5	1918	80.6	1936	79.7	1954	4.4	1972	68.9	1990	142.6
1901	2.7	1919	63.6	1937	114.4	1955	38	1973	38	1991	145.7
1902	5	1920	37.6	1938	109.6	1956	141.7	1974	34.5	1992	94.3
1903	24.4	1921	26.1	1939	88.8	1957	190.2	1975	15.5	1993	54.6
1904	42	1922	14.2	1940	67.8	1958	184.8	1976	12.6	1994	29.9
1905	63.5	1923	5.8	1941	47.5	1959	159	1977	27.5	1995	17.5
1906	53.8	1924	16.7	1942	30.6	1960	112.3	1978	92.5	1996	8.6
1907	62	1925	44.3	1943	16.3	1961	53.9	1979	155.4	1997	21.5
1908	48.5	1926	63.9	1944	9.6	1962	37.6	1980	154.6	1998	64.3
1909	43.9	1927	69	1945	33.2	1963	27.9	1981	140.4	1999	93.3
1910	18.6	1928	77.8	1946	92.6	1964	10.2	1982	115.9	2000	199.6
1911	5.7	1929	64.9	1947	151.6	1965	15.1	1983	66.6	2001	111
1912	3.6	1930	35.7	1948	136.3	1966	47	1984	45.9	2002	104
1913	1.4	1931	21.2	1949	134.7	1967	93.8	1985	17.9	2003	63.7
1914	9.6	1932	11.1	1950	83.9	1968	105.9	1986	13.4	2004	40.4
1915	47.4	1933	5.7	1951	69.4	1969	105.5	1987	29.4	2005	29.8
1916	57.1	1934	8.7	1952	31.5	1970	104.5	1988	100.2	2006	15.2
1917	103.9	1935	36.1	1953	13.9	1971	66.6	1989	157.6	2007	7.55